Predicting Recessions:
Forecasting US GDP Growth through Supervised Learning

Russ Islam
Department of Electrical Engineering, Stanford University

Abstract

Machine learning algorithms have gained much popularity in finance, where the abundance of training examples and high-frequency sampling rates produce datasets that are amenable to successful regression. In macroeconomics, however, where data is scarce and sampling rates are far lower, learning algorithms have not been extensively explored, and even within the sparse literature success has been limited. In this paper I explore the ability of Support Vector Regression (SVR) and feedforward neural networks to predict the 2008 Great Recession when trained only on macroeconomic data from 1978 to 2005. Their performance is compared to that of a standard ARIMA model taken from recent economic literature.

I. Background and Formulation

The motivation behind this research is two-fold: to estimate the input-output relationship between various macroeconomic indicators and US quarterly GDP, and to select which indicators improve this estimation. Leading economic indicators (LEIs), such as the value of the S&P 500, are time series that are believed to serve as bellwethers of the future economy. To test this conjecture, I consider the following 5 different LEIs for their usefulness in predicting quarterly GDP (these constitute 5 of the 10 indicators used in The Conference Board's Leading Economic Index):

1. Value of the S&P 500
2. Number of weekly initial jobless claims
3. Velocity of M2 money supply
4. Number of new private building permits
5. The Consumer Sentiment Index, published by the University of Michigan

In addition to the quarterly GDP time series itself, each of these time series is averaged quarterly. Because the Consumer Sentiment Index was first published in 1978, we consider data only from 1978 onwards for all five indicators and for GDP. Limiting our dataset to this time period is justified for economic reasons as well: 1973 was the year of publication of the Black-Scholes equation, which led to a dramatic surge in the activity of the Chicago Board of Options Exchange (CBOE), and in the same year most of the world's currencies switched from fixed to floating exchange regimes. Under the assumption that the structure of the US economy (and the global economy) underwent a significant change during the mid-1970s, it is justified to restrict our input time series to the period after 1978.

This time series forecasting problem may be formalized by considering the general Nonlinear AutoRegressive eXogenous (NARX) model [1], which relates the current value of an output time series, y_t, to previous values of that time series and to current and previous values of explanatory input time series (for instance, a and b):

$$y_t = F(y_{t-1}, y_{t-2}, \dots, a_t, a_{t-1}, a_{t-2}, \dots, b_t, b_{t-1}, b_{t-2}, \dots)$$

I employ SVRs and neural networks to estimate the (possibly) nonlinear function F and to select useful exogenous inputs from the aforementioned 5 LEIs. I add a further restriction to the NARX model by excluding current values of the exogenous inputs; this allows us to predict future values of GDP without knowing the future values of the inputs. One-quarter-lagged and two-quarter-lagged versions of all 5 indicators as well as quarterly GDP comprise our dataset of 12 time series.

II. Feature Selection and SVRs

Forward search feature selection was conducted to choose indicator time series that reduce the generalization root mean square error (RMSE) of epsilon-SVR when trained on data from 1978:1 to 2005:4 and tested on data from 2006:1 to 2013:3. The SVR was implemented using both a linear kernel and a second order polynomial kernel; forward search was conducted several times on various values of the SVR cost parameter *C* and the insensitivity parameter epsilon to find their optimal values. This process was repeated for one-quarter-ahead and three-quarter-ahead prediction, the results of which are shown in Table 1 and Table 2, respectively.

Plots on the following page overlay the actual GDP values, the predictions of the SVRs, and the predictions of an ARIMA(1,1,1) model. The parameters for the ARIMA model are taken from [2], whose authors argue that these are the optimal values for GDP prediction. Compared to those of ARIMA, the linear SVR predictions are clearly closer to the actual values for most quarters. The ARIMA model overestimates GDP growth immediately before the

Table 1: One-step-ahead forecasting results for forward search using epsilon-SVR. Features are numbered according to the list given on the previous page. RMSE is given in units of billions of chained 2009 US dollars. The initial feature set contains GDP lag1.

One-step-ahead forecasting	Order of Poly Kernel	
	1 (Linear)	2
Optimal *C*	2	10
Optimal epsilon	90	630
RMSE	89.6	152.8
Selected Features in Descending Order of Selection	5 (lag 1) 1 (lag 2) GDP (lag 2) 3 (lag 2)	4 (lag 1) 3 (lag 1)

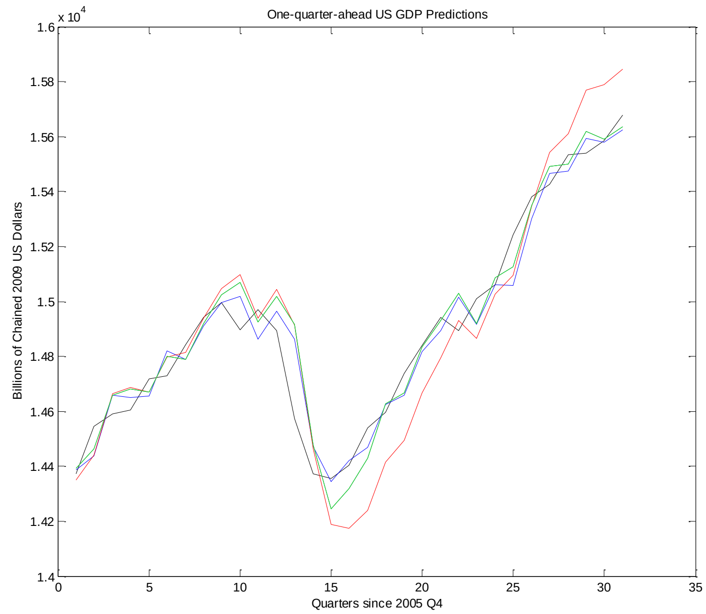
Table 2: Three-steps-ahead forecasting results for forward search using epsilon-SVR. Features are numbered according to the list given on the previous page. RMSE is given in units of billions of chained 2009 US dollars. The initial feature set contains GDP lag1.

Three-steps-ahead forecasting	Order of Poly Kernel	
	1 (Linear)	2
Optimal *C*	1.3	10
Optimal epsilon	60	500
RMSE	210.7	270.8
Selected Features in Descending Order of Selection	5 (lag 1) 4 (lag 1) 4 (lag 2) 1 (lag 2) 5 (lag 2)	GDP (lag 2)

downturn and is overly pessimistic when predicting the severity of the recession. Nevertheless, there is a delayed response apparent in both SVR and ARIMA, though SVR appears to have a greater ability to correct itself.

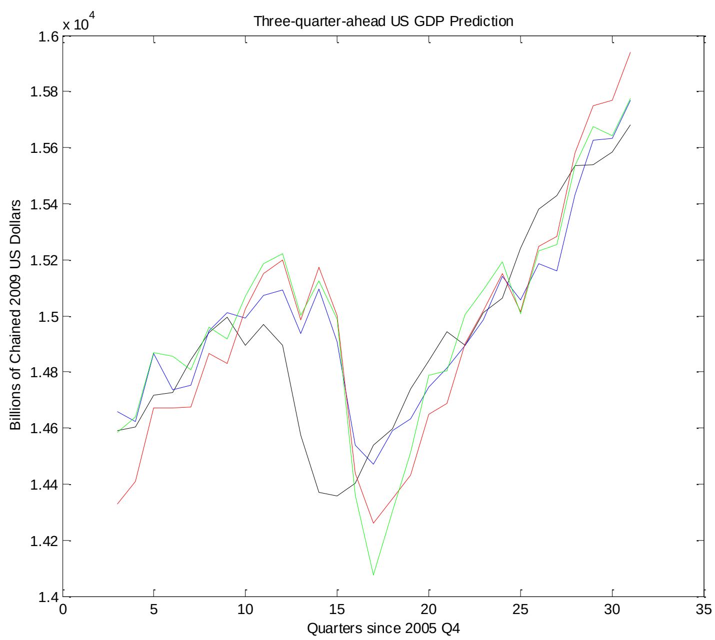
Plot 1:
One-step-ahead
GDP Prediction

Actual GDP value - black
SVM (linear) - blue
SVM (order 2 poly) – red
ARIMA(1,1,1) - green



Plot 2:
Three-steps-ahead
GDP Prediction

Actual GDP value - black
SVM (linear) - blue
SVM (order 2 poly) – red
ARIMA(1,1,1) - green



It is interesting to note that fewer features are selected when the order of the polynomial kernel increases from 1 to 2. Among the indicators, the Consumer Sentiment Index is especially popular, as is the number of building permits. The two-step-lagged S&P 500 is also seen to be chosen by the linear SVRs. Somewhat surprisingly, some indicators are selected even before two-step-lagged GDP is. In none of the simulations does feature 2 (initial jobless claims) improve SVR performance.

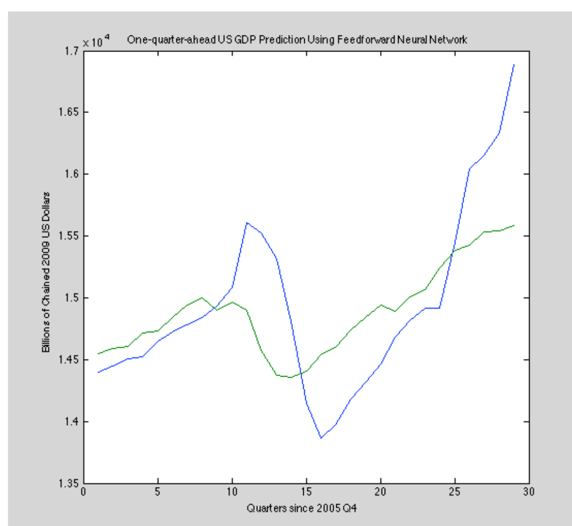
These results demonstrate that SVR can make use of relationships between LEIs to produce models with predictive power greater than that of the simple ARIMA models that have been so popular among econometricians. LEIs indeed provide some information about the future state of the economy and merit additional quantitative research.

III. Neural Networks

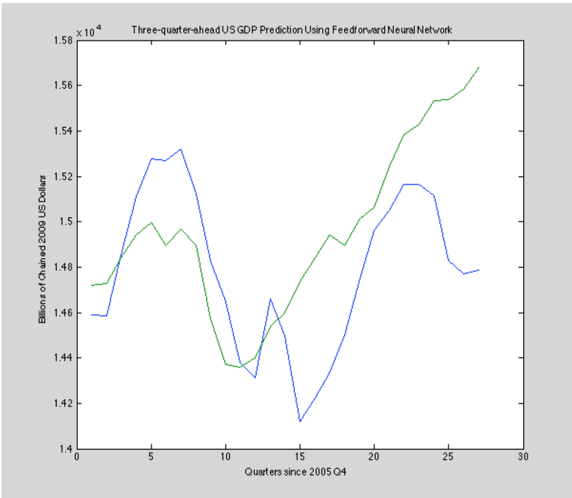
A feedforward neural network with one hidden layer and [1:2] delays for both the inputs (all 5 LEIs) and the output was trained and tested on the same dataset using the Bayesian regulation learning rule (the network was trained in open configuration and tested in closed feedback configuration). The performance of the network averaged over 10 trials was measured with the number of neurons in the hidden layer varied from 5 to 36 to find the optimal number of neurons. The small training set (approximately 100 samples) appears to make the performance of the neural network heavily dependent on the initial randomization of its weights, so the optimal number of neurons found varied between 6 and 13 each time the optimization was run. A network with 10 neurons in the hidden layer was used to produce the plots that follow.

With this dataset, neural networks exhibit significantly less predictive power than either SVRs or ARIMA. The networks overfit the training set and display bias problems that prevent accurate generalization to the test set. Even after optimizing the number of neurons in the hidden layer the generalization error remains high. Although in general neural networks are well-suited to estimate the solutions to time series problems, the sparseness of this dataset causes the behavior of neural networks to be largely erratic and substantially influenced by initial randomization. Even if neural networks can theoretically provide better fits than SVR or ARIMA, the reproducibility of their behavior is harmed by the small size of the training set. The authors of [3], who explored the application of neural networks to predict the GDP of Canada, also note that small training sets deteriorate the performance of neural networks, and they surmise that at least 300 observations are needed to achieve results better than those of traditional linear regression models.

Plot 3: Feedfoward neural network for one-step-ahead prediction.



Plot 4: Feedforward neural network for three-step-ahead prediction.



IV. Conclusions and Further Work

With only a small training set, Support Vector Regression with a linear kernel generalizes well to its test set and displays significantly better prediction performance than does ARIMA(1,1,1), though feedforward neural networks with a single hidden layer suffer from bias problems and perform worse than both SVR and ARIMA. The results of this paper indicate that SVR has the potential to become a useful tool for macroeconomic prediction and for regression in other fields with scarce data and/or low sampling rates. Compared to ARIMA, SVR exhibits less erroneous variance during testing and is able to correct itself more robustly when it veers too far from true labels. Forward search with SVR also discovered the predictive usefulness of different LEIs. On the other hand, though neural networks perform poorly with exceptionally small training sets, one should not conclude that they are not useful in low-sampling-rate problems. Simulations with larger training sets could produce more fruitful results.

References

[1] R. Salat, et al., "Black-Box system identification by means of Support Vector Regression and Imperialist Competitive Algorithm," in *Przeglad Elektrotechniczny*, September 2013.

[2] E. Andrei and E. Bugudui, "Econometric Modeling of GDP Time Series," in *Theoretical and Applied Economics*, vol. XVIII (2011), No. 10(563), pp. 91-98.

[3] G. Tkacz and S. Hu, "Forecasting GDP Growth Using Artificial Neural Networks", Bank of Canada working paper 99-3 (http://www.bankofcanada.ca/wp-content/uploads/2010/05/wp99-3.pdf).

The dataset was obtained from the Federal Reserve Economic Data (FRED) website of the Federal Reserve Bank of St. Louis (http://research.stlouisfed.org/fred2/).